Molecular Simulation on the K Computer and towards Fugaku

The 2nd R-CCS International Symposium
14:45–15:00, 17 February 2020
Nichii Gakkan Kobe Port Island Center, Kobe

RIKEN Center for Computational Science, Kobe, Japan

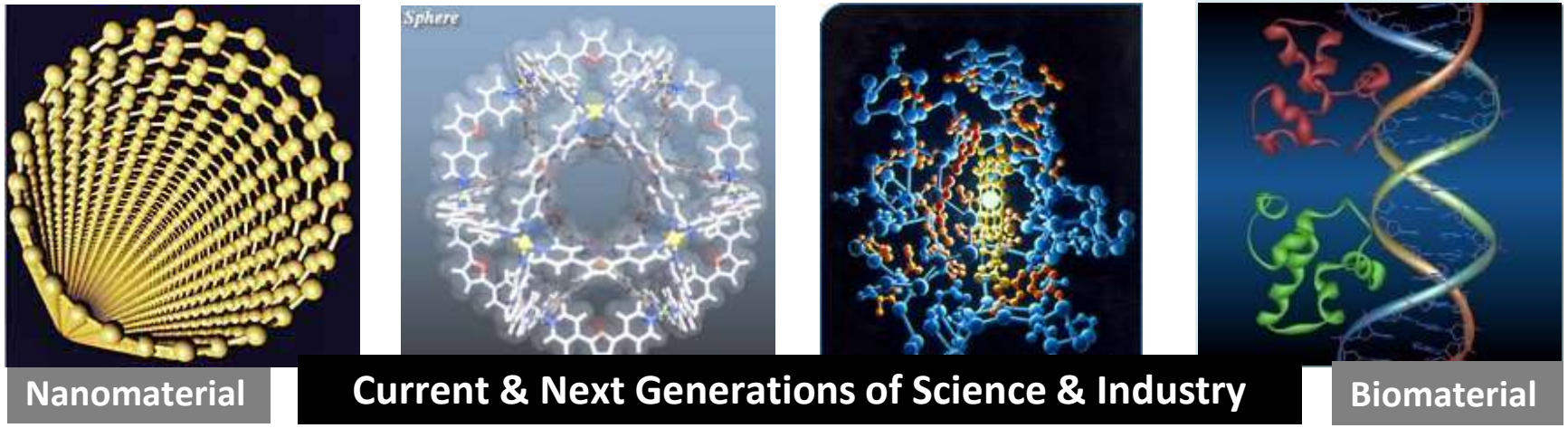
Takahito Nakajima

Computational Molecular Science Research Team

- The main interest in our team is the development of theoretical and computational molecular science based on quantum chemistry. Our main background is molecular orbital (MO) theory, for which the many-electron Schrödinger equation is approximately solved by the 1-electron equation.
- In addition, the target is to develop interdisciplinary research by the combination of quantum chemistry with other research fields such as condensed matter physics, biological chemistry, computer science, informatics, data science, ….

Theoretical & Computational Molecular Science

- With the emergence of peta- or exa-scale computing platforms, we are entering a new period of modeling. The molecular simulations can be carried out for larger, more complicated, and more realistic systems than ever before.



- To benefit from K/Fugaku exhaustively, we are forging **novel molecular science based on original theorizing** and **K/Fugaku-compatible software** that can predict outcomes of molecular experiments.

Our Research Goals

To lead the way toward a new frontier of theoretical and computational molecular science, the project involves the novel development of theory, algorithm, and software, which will be realized through collaborative use of K and Fugaku across the fields of computational science and computer science.

1. **Innovation of Theoretical and Computational Molecular Science Based on Next-Generation Molecular Theory**

   to realize an improved and updated theoretical molecular science by developing our original theorizing to handle large and complicated molecules with high accuracy

2. **Development of New Quantum Chemistry Software "NTChem"**

   to provide the users to a high-performance software package for molecular electronic-structure calculations for general purpose

3. **Establishment of Materials Informatics on Massively Parallel Supercomputers**

   to develop efficient schemes for materials informatics with high-throughput calculations on K/Fugaku

Our Research Goals

To lead the way toward a new frontier of theoretical and computational molecular science, the project involves the novel development of theory, algorithm, and software, which will be realized through collaborative use of K and Fugaku across the fields of computational science and computer science.

1. **Innovation of Theoretical and Computational Molecular Science Based on Next-Generation Molecular Theory**

   to realize an improved and updated theoretical molecular science by developing our original theorizing to handle large and complicated molecules with high accuracy

2. **Development of New Quantum Chemistry Software "NTChem"**

   to provide the users to a high-performance software package for molecular electronic-structure calculations for general purpose

3. **Establishment of Materials Informatics on Massively Parallel Supercomputers**

   to develop efficient schemes for materials informatics with high-throughput calculations on K/Fugaku

NTChem

- QC software comprises immensely useful tools in material and biological science researches.
- The development of QC software in Japan falls far short of that in Western communities.

World-Wide QC Programs

-- Gaussian (Gaussian, Inc.)
-- GAMESS (Iowa State Univ.)
-- Molpro (Stuttgart, Cardiff)
-- Molcas (Lund Univ.)
-- NWChem (PNNL)
-- Q-Chem (Q-Chem, Inc.)
-- TURBOMOLE (GmbH)

ORCA
DALTON
GAMESS
NWChem
High Performance Computational Chemistry Software
Q-CHEM

- In the K project, we decided to develop a comprehensive new program suite of *ab initio* QC made in Japan from scratch.

NTChem

Features of NTChem

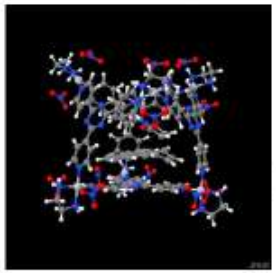
NTChem implements not only standard QC approaches, but also our-own original and improved QC methods that we have developed in our research works.

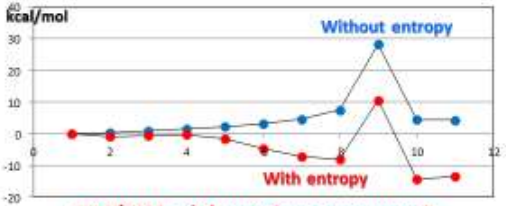
- Electronic structure calc. of the ground state of molecules based on DFT&HF.
- Linear- or low-scaling DFT: RI-DFT, PS-DFT, GAPW, GFC, dual-level DFT.
- Low-scaling SCF calculation using diagonalization-free approaches.
- Excited-state DFT calculation: TDDFT, DFT-TP.
- Accurate correlation methods for ground and excited states: MP2, CC, QMC.
- Relativistic electronic structure calc. with SO interactions: DK, RA, RESC.
- Model calculations for large molecules: QM/MM, ONIOM, ADMA, MTA.
- Calculation of solvent effects: COSMO, ASEP/MD, QM/MM-MD.
- Efficient calculation for chemical reaction pathway: NEB, String.
- Ab initio molecular dynamics calculation.
- Calculation of electric and magnetic properties: Polarizability, NMR, EPR.
- Population analysis: Mulliken, Löwdin, NBO.
- Orbital interaction analysis: MIO, PIO.
- Massively parallel computing on K and PC architectures: HF, DFT, MP2, QMC.

A Wide Variety of Applications with NTChem

Diels–Alder Reaction in Supramolecule

Free Energy Path of Diels–Alder Reaction in Cage



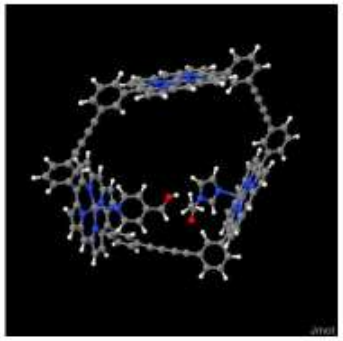


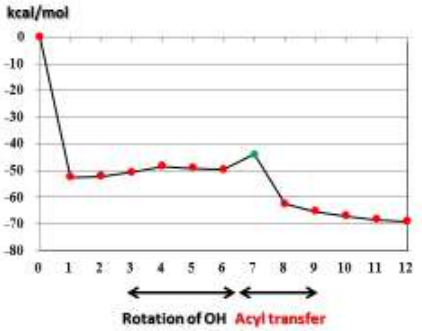
QM/MM minimum free energy path

$$A(\bar{\mathbf{R}}^{\mathrm{G}},\mathbf{R}^{\mathrm{II}}) = -\frac{1}{\beta}\ln\int d\mathbf{R}^{\mathrm{II}}\exp\left(-\beta E(\bar{\mathbf{R}}^{\mathrm{G}},\mathbf{R}^{\mathrm{II}})\right)$$

- When no entropic contribution is considered, the activation energy with the cage (28kcal/mol) is comparable to that without the cage (23kcal/mol).
- On the other hand, the activation energy with the entropic contribution (11kcal/mol) is much lower than that without it.
- The role of the cage in this reaction is not to accelerate the reaction favorably in enthalpy, but to accelerate the reaction in entropy by changing the vibration of the cage into the heat.

Catalytic Acyl Transfer by Cyclic Porphyrin Trimer

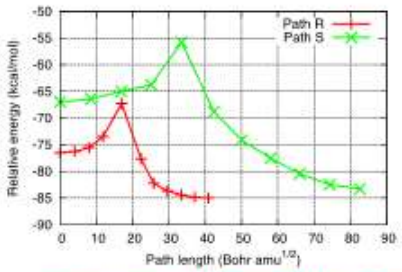




- The acyl transfer reaction in the cyclic Zn porphyrin trimer proceeds favorably with a low activation energy (9 kcal/mol) and a large reaction heat (–17 kcal/mol) in contrast to that without the trimer.
- The reaction can easily progress by flexible change of the cage configuration.

Enantioselectivity in Cage Complex

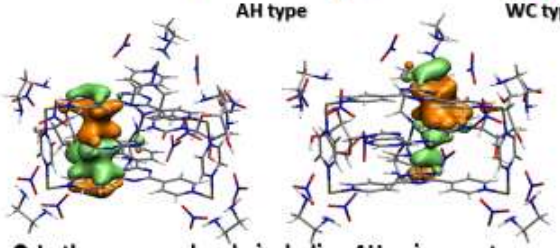
Minimum Energy Path of Aza-Cope Rearrangement in Cage



- Encapsulation in the cage stabilizes prochiral substrates in the initial stage of the reaction. The prochiral substrate for R enantiomer is more stabilized than that for S one.
- The whole energy path for R enantiomer is also more stabilized than that for S one.
- The high enantioselectivity is ascribed to the difference in the stability of prochiral substrate encapsulated in the host cage.

Ootani, Akinaga, Nakajima, J. Comput. Chem. 36, 459 (2015).

Selective Recognition in Supramolecule



MIOs which represent the first & second strongest interactions between the cage and the AH pair

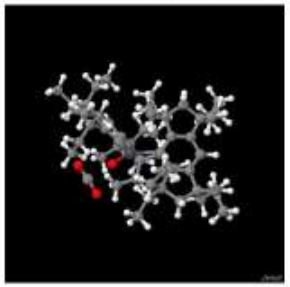
- In the supramolecule including AH pair, π systems around N atom in A and O atom in U are strongly bound to two aromatic π systems in panels of the cage. This supports that AH base pair is more stable than WC one within the cage.

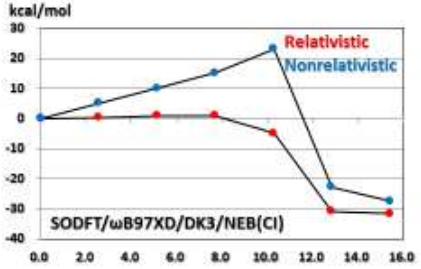
A Wide Variety of Applications with NTChem

Chemical Reactivity with Relativity

Insertion Reaction of CO_2 to Pb=O Bond in Plumbanone

- We performed the predictive simulation in which Ge was substituted by Pb.



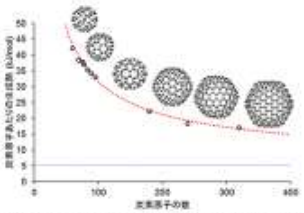


- QC simulation with relativistic effects predicts that plumbanone shows high catalytic activity for CO_2 insertion as well as germanone.
- Relativistic effects are important. The reaction cannot proceed without relativistic effects.

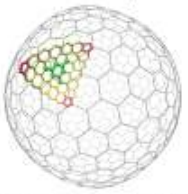
This kind of calculation cannot be performed unless NTChem is used.

K Provides New Insights into Fullerenes

Recently, we have made use of the tremendous power of the K computer, together with NTChem, to provide new insights into the properties of fullerenes. We can predict, with high accuracy, the property known as the heat of formation, which is important for comprehending how readily a compound will form. In our investigation, we have made a very precise prediction for not only C_{60}, which is the most common fullerene, but also a series of larger fullerenes, including C_{360}. The results of this project suggest an important contributing factor for the unique properties of fullerenes and other types of nanomaterials. We find that the heats of formation for fullerenes become closer to that of graphene as the number of carbon atoms increases, but only at a slow pace. Our finding suggests that new interesting properties that show the character of both fullerenes and graphenes may appear if much larger fullerenes are synthesized.



Calculated heats of formation per carbon (kJ mol^{-1})



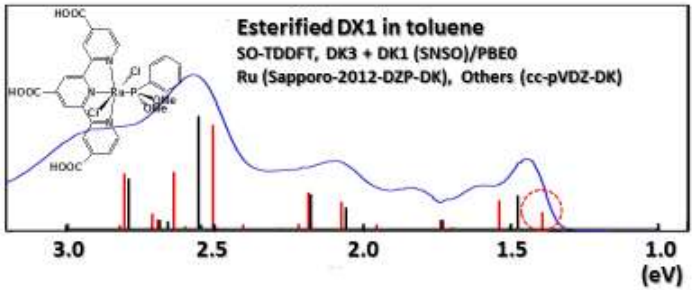
The red sections are high-strain pentagon "vertices" that differentiate fullerenes from graphene.

【Press release】 http://www.riken.jp/pr/press/2016/20160316_1/

Chan, Kawashima, Katouda, Nakajima, Hirao, JACS, 138, 1420 (2016).

Absorption Spectra of DX1

To examine spin-forbidden transitions of DX1 in details, transition energies for DX1 were calculated by two-component relativistic SO-TDDFT.

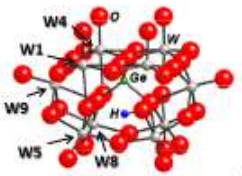


- SO-TDDFT can reproduce experimental findings.
- The SO effect is large on the first peak around 1.35 eV, whereas the effect is relatively small on other peaks. The first peak shifts to the longer wavelength side due to strong SO interaction.

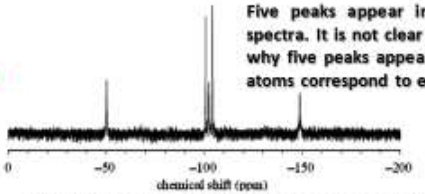
Sugahara, Satake, Kamata, Nakajima, Mizuno, Angew. Chem. Int. Ed, 53, 13464 (2014).

NMR Chemical Shifts

^{183}W NMR Chemical Shifts in POM



$[\gamma\text{-HGeW}_{10}O_{36}]^{7-}$
(1プロトン和体)



Five peaks appear in experimental spectra. It is not clear experimentally why five peaks appear and which W atoms correspond to each peak.

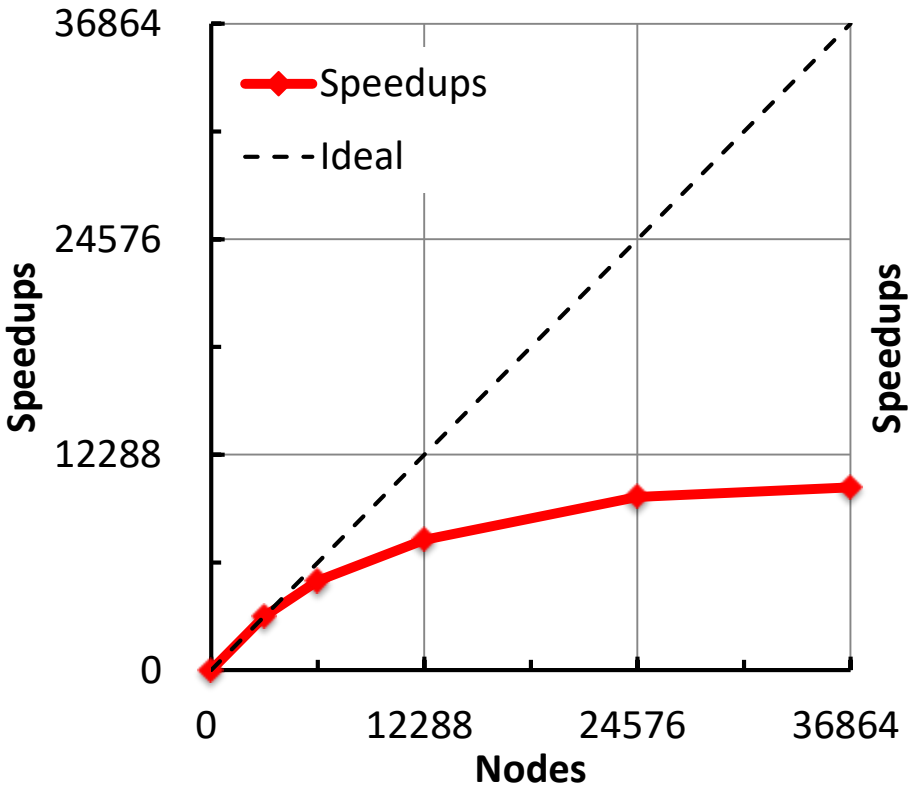
^{183}W NMR Spectra of −7-charged germanodecatungstate

	spin-free IGLO (DMSO)		SO-IGLO (gas)		SO-IGLO (DMSO)		Assign (ppm)
	Shild. const. (ppm)	Chemical shift (ppm)	Shild. const. (ppm)	Chemical shift (ppm)	Shild. const. (ppm)	Chemical shift (ppm)	
W1	1013.0	-411.5	1014.3	-301.8	913.4	-85.1	-103.0
W4	1027.8	-426.3	1016.4	-303.9	911.1	-82.7	-102.0
W5	851.5	-250.0	917.1	-204.6	876.5	-48.2	-51.0
W8	953.1	-351.7	961.5	-249.0	917.2	-88.8	-105.0
W9	1009.8	-408.4	1085.7	-373.3	1032.1	-203.8	-150.0

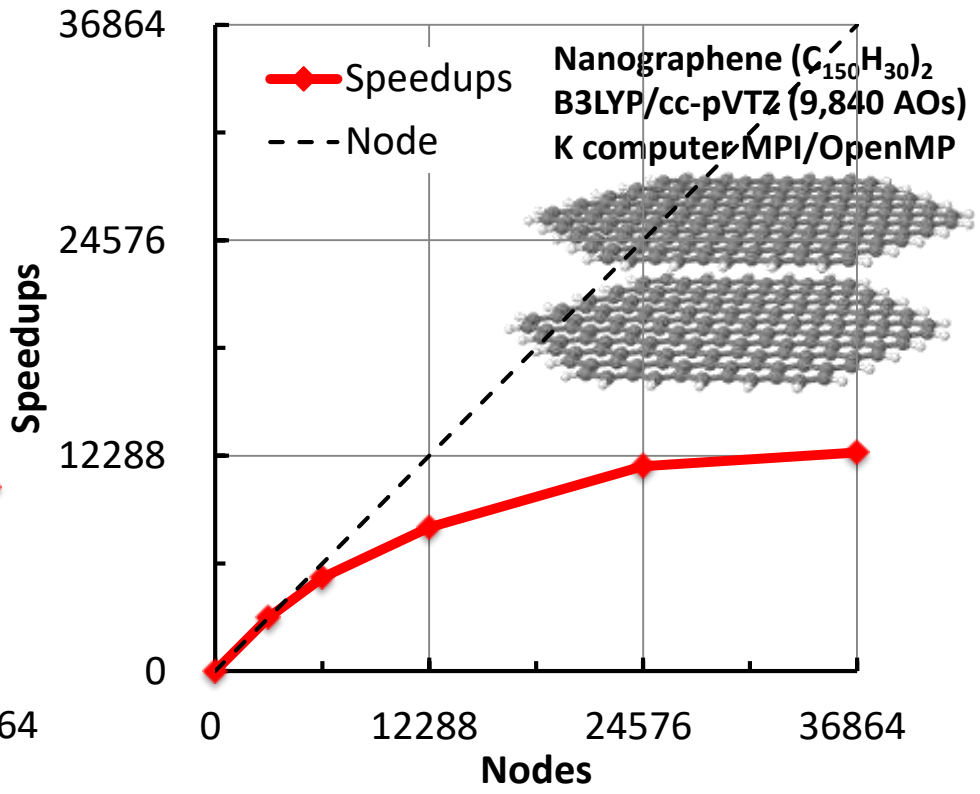
Assignment of ^{183}W NMR spectra of POM was realized by cooperation with theoretical and experimental studies.

Performance of Parallel Computing with NTChem

- NTChem is designed for high-performance calculations on the computers with several tens of thousands of nodes for DFT, TDDFT, MP2, and QMC.



Total SCF (Fock matrix calc. + diagonalization)

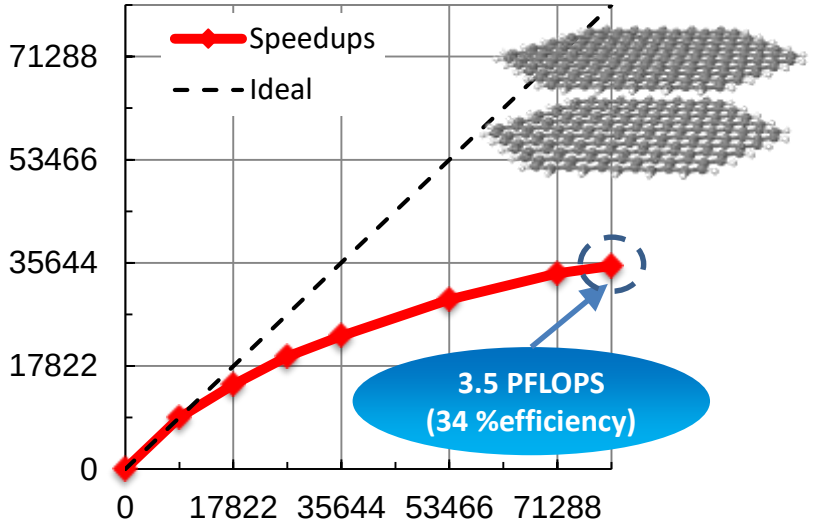


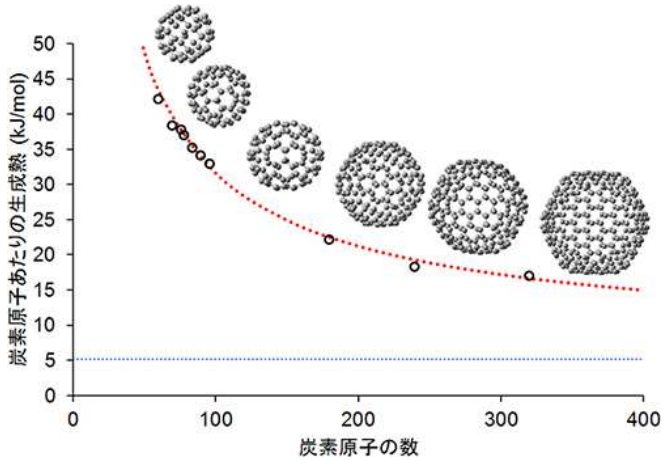
Fock matrix calculation

Parallel scalability is kept up to 36,864 nodes of K: Elapsed time 95 min using 36,864 nodes

Hybrid Parallel MP2 Algorithm in NTChem

- We have proposed an efficient MPI/OpenMP hybrid parallel algorithm of the MP2 electron-correlation calculation.





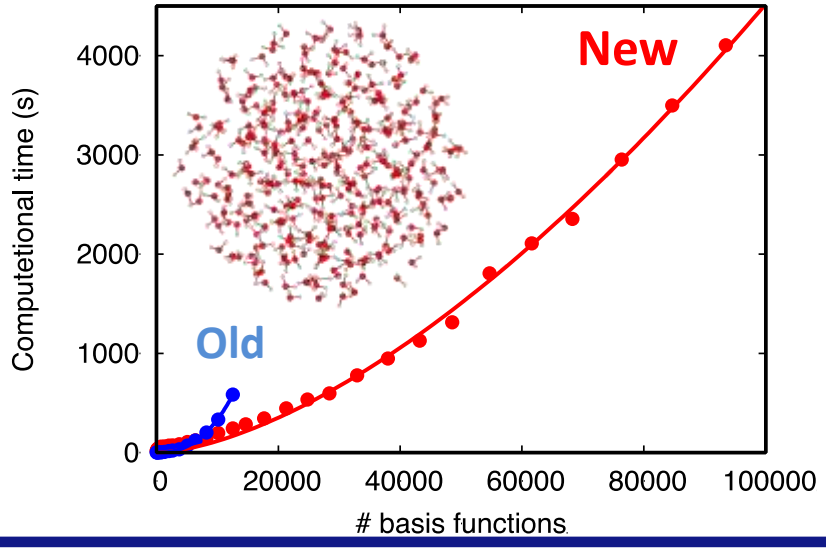
Calculated heats of formation per carbon (kJ mol^{-1})

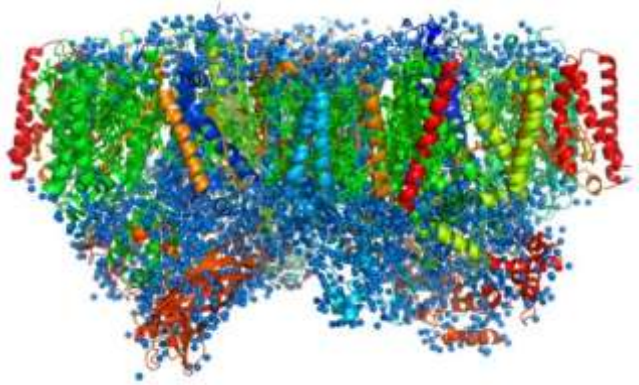
- The good parallel scaling is achieved with our new algorithm on "K". Elapsed time is 11 mins using 80,199 nodes of "K" for nanographene $(C_{150}H_{30})_2$. The world's largest wavefunction-based electron-correlation calculation (Left Figure)
- By using the same MP2 algorithm on "K", we can predict, with high accuracy, the property known as the heat of formation of large fullerenes. Our finding suggests that new interesting properties that show the character of both fullerenes and graphenes appear if much larger fullerenes are synthesized. (Right Figure)

Press release: http://www.riken.jp/pr/press/2016/20160316_1/

Development of NTChem

- We are currently developing a new version of NTChem. It adopts a memory-distributed sparse-matrix algorithm based on NTPoly to overcome available memory sizes per a computer node. It also adopts an algorithm with regular communication patterns that map to high dimensional torus network architectures to manage significant irregular data movement.
- We hope that *ab initio* calculations for large molecules including 1M atoms will be realized on Fugaku because the new version is capable of calculating molecules with 10K atoms on "K" in the current stage.
- In future, we will perform novel computational applications for realistic molecular systems such as whole photosystem II (1.2M atoms) on Fugaku.





NTChem

Our Research Goals

To lead the way toward a new frontier of theoretical and computational molecular science, the project involves the novel development of theory, algorithm, and software, which will be realized through collaborative use of K and Fugaku across the fields of computational science and computer science.

1. Innovation of Theoretical and Computational Molecular Science Based on Next-Generation Molecular Theory

   to realize an improved and updated theoretical molecular science by developing our original theorizing to handle large and complicated molecules with high accuracy

2. Development of New Quantum Chemistry Software “NTChem”

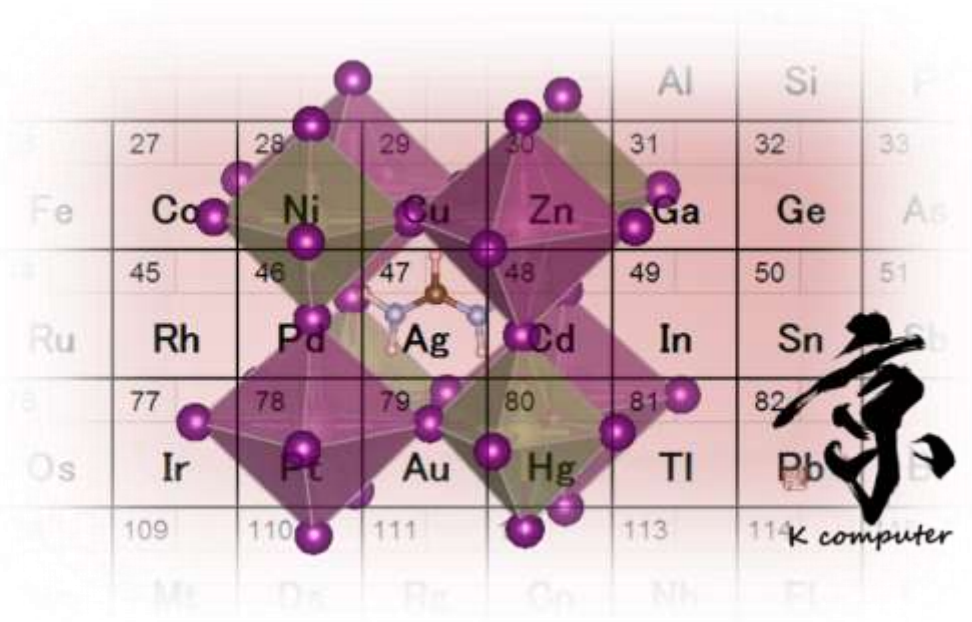
   to provide the users to a high-performance software package for molecular electronic-structure calculations for general purpose

3. **Establishment of Materials Informatics on Massively Parallel Supercomputers**

   to develop efficient schemes for materials informatics with high-throughput calculations on K/Fugaku

Materials Design for Solar Cells

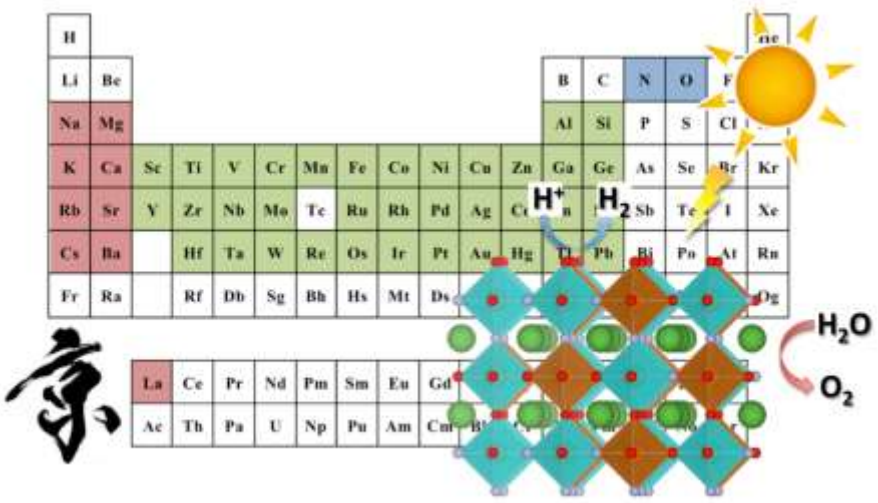
- To search for **novel Pb-free perovskite solar cells** containing non-toxic and widely available metals, we performed a systematic high-throughput first-principles simulation on **the K computer** for **11,025 compositions of perovskites** in ABX_3 and $A_2BB'X_6$ forms, where A is an organic or inorganic component, B/B' is a metal atom, and X is a halogen atom.
- By applying the screening procedure to all of the computed compounds, we discovered novel candidates for environmentally friendly Pb-free perovskite solar cells and proposed **51 low-toxic halide perovskites**, most of which are proposed newly in our work.
- The potential candidates are categorized as **only 6 types** based on the combination of groups to which two metal elements (at B- and B'-sites) belong in the periodic table; **group 14–14**, **group 13–15**, **group 11–11**, **group 9–13**, **group 11–13**, **group 11–15**.



Press release: http://www.riken.jp/pr/press/2017/20171005_1/

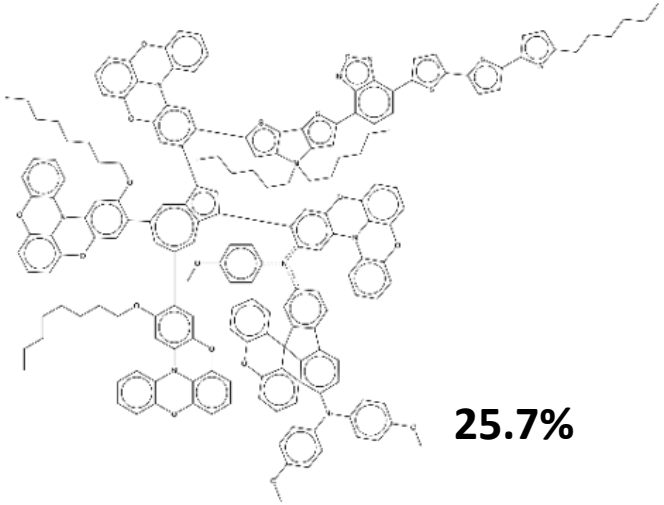
Materials Design for Photocatalysts

- We performed a large-scale computational screening for 29,160 perovskite oxynitrides and oxides to explore novel visible-light water-splitting photocatalysts. We can propose 42 potential candidates, most of which are newly discovered in our study.
- We can suggest that the perovskite oxynitrides with d^0/d^{10} electronic configuration and the perovskite oxides with partially occupied d electrons are desirable for the visible-light driven water-splitting photocatalysts with adequate band gaps and band edge positions.
- By selecting the perovskites containing cheaper elements than vanadium from the 42 potential candidates, we can propose novel 6 compounds, $NaWO_2N$, KWO_2N, $MgWON_2$, $CaVO_2N$, $CaAl_{1/3}W_{2/3}O_2N$, and $CaV_{2/3}Fe_{1/3}O_3$, as rare-metal-free photocatalysts.

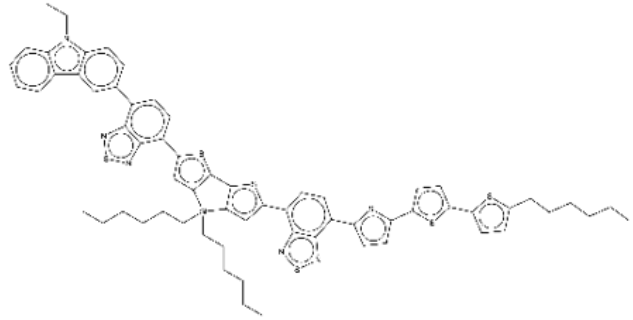


Materials-Informatics Driven Design of HTMs

- We proposed new hole-transport materials (HTMs) by using the materials informatics and the simulations with NTChem on "K". We firstly created a machine-learning model to predict the power conversion efficiency of perovskite solar cells (PSCs) for the molecules generated from structural combinations of the fragments of known HTMs. The quantum-chemical electronic descriptors were evaluated by NTChem on "K". Then, to select HTM with the highest PCE, we adopted the particle swarm optimization (PSO) with the GPR model. In details, the optimal set of fragments for HTM with the highest PCE was selected by PSO so as to optimize the acquisition function obtained from the GPR model. We achieved the efficient search for novel candidates for HTMs of PSCs from the vast chemical space (32,294,400 molecules).



25.7%



24.7%

Candidate HTMs predicted by materials informatics

Summary & Future Plans

- To lead the way toward a new frontier of theoretical and computational molecular science, we are developing novel theory, algorithm, and software, which will be realized through collaborative use of K and Fugaku across the fields of computational science and computer science.
- NTChem is a high-performance software package for molecular electronic structure calculation for general purpose, and is a useful tool in various computational studies for large and complicated molecular systems.
- NTChem is a community-based software package. We intend to continue adopting users' requests with the aim of making the program more convenient and user-friendly for researchers in various research fields.

  User cases

  ・Toyobo ・Mitsubishi Chem. ・RIKEN ・Hokkaido U. ・U. Tokyo ・Kyoto U. ・Kobe U. ・Yokohama-City U. ・Gifu U. ・JAEA ・Open to the public on K and SC centers (IMS, FOCUS, Nagoya U., Kyusyu U., Tokyo Tech.)

http://www.r-ccs.riken.jp/software_center/software/ntchem/

Summary & Future Plans

- We plan to realize accelerating the discovery of novel energy materials such as photovoltaics, photocatalysts, and polymer batteries by establishing new materials informatics techniques that assimilate large databases and mathematics in materials science with high-throughput calculations on Fugaku. We currently have several plans to collaborate with experimental researchers and industries in order to discover new materials that are highly efficient, low cost, environmentally clean, and sustainable. We would like to realize the novel material designs by the co-design with experiments, theorizing, simulation, and informatics.